



Bruedern, J., & Wooley, T. D. (2016). The Hasse principle for systems of diagonal cubic forms. *Mathematische Annalen*, 364(3-4), 1255-1274. <https://doi.org/10.1007/s00208-015-1249-1>

Peer reviewed version

Link to published version (if available):  
[10.1007/s00208-015-1249-1](https://doi.org/10.1007/s00208-015-1249-1)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

The final publication is available at Springer via <http://dx.doi.org/10.1007/s00208-015-1249-1>

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# THE HASSE PRINCIPLE FOR SYSTEMS OF DIAGONAL CUBIC FORMS

JÖRG BRÜDERN AND TREVOR D. WOOLEY

ABSTRACT. We establish the Hasse Principle for systems of  $r$  simultaneous diagonal cubic equations whenever the number of variables exceeds  $6r$  and the associated coefficient matrix contains no singular  $r \times r$  submatrix, thereby achieving the theoretical limit of the circle method for such systems.

## 1. INTRODUCTION

The Diophantine analysis of systems of diagonal equations was pioneered by Davenport and Lewis with a pivotal contribution on pairs of cubic forms [7], followed by work on more general systems [8]. For natural numbers  $r, s$  and an  $r \times s$  integral matrix  $(c_{ij})$ , they applied the circle method to the system

$$\sum_{j=1}^s c_{ij} x_j^3 = 0 \quad (1 \leq i \leq r), \quad (1.1)$$

and when  $s \geq 27r^2 \log 9r$  were able to show that (1.1) has infinitely many primitive integral solutions. Even a casual practitioner in the field will acknowledge that the implicit use of mean values demands at least  $6r + 1$  variables in the system for the circle method to be applicable. We now attain this theoretical limit, surmounting the obstacles encountered by previous writers.

**Theorem 1.1.** *Let  $s > 6r$  and suppose that the matrix  $(c_{ij})$  contains no singular  $r \times r$  submatrix. Then, whenever the system (1.1) has non-zero  $p$ -adic solutions for all primes  $p$ , it has infinitely many primitive integral solutions.*

The conclusion of Theorem 1.1 may be interpreted as a Hasse principle for systems of diagonal cubic forms in general position. As we remark in §4, the condition on the matrix of coefficients can be relaxed considerably. Should the local solubility conditions be met, our methods show that the number  $N(P)$  of integral solutions of (1.1) with  $\mathbf{x} \in [-P, P]^s$  satisfies  $N(P) \gg P^{s-3r}$ . We note that work of the first author joint with Atkinson and Cook [1] implies that for  $p > 9^{r+1}$  the  $p$ -adic solubility hypothesis in Theorem 1.1 is void.

Early work on this subject concentrated on methods designed to disentangle the system so as to invoke results on single equations. The most recent such

---

2010 *Mathematics Subject Classification.* 11D72, 11P55, 11E76.

*Key words and phrases.* Cubic Diophantine equations, Hardy-Littlewood method.

The authors are grateful to the Hausdorff Research Institute for Mathematics in Bonn for excellent working conditions that made the writing of this paper feasible. The support of the Akademie der Wissenschaften zu Göttingen is also gratefully acknowledged.

contribution is Brüdern and Cook [3] where the condition  $s > 7r$  is imposed on the number of variables. Such methods are incapable of establishing the conclusion of Theorem 1.1 unless one is prepared to invoke conditional mean value estimates that depend on speculative Riemann hypotheses for global Hasse-Weil  $L$ -functions (see Hooley [11, 12] and Heath-Brown [10]).

When  $r = 1$ , the conclusion of Theorem 1.1 is due to Baker [2]. For  $r \geq 2$ , the present authors [4] identified features of fully entangled systems of equations which permit highly efficient use of divisor estimates in bounding associated multidimensional mean values. These allow treatment of systems in  $6r + 3$  variables. By a method special to the case  $r = 2$ , we established that case of Theorem 1.1 in more general form (see [5]). In this paper we instead develop a recursive process that relates mean values associated with the original system to a one-dimensional sixth moment of a smooth Weyl sum on the one hand, and on the other to another system of the shape (1.1), but of much larger format. The new system is designed in such a way that the methods of [4] provide very nearly square-root cancellation. By comparison with older routines, we are forced to incorporate the losses implied by the use of a sixth moment of a smooth Weyl sum only once, as opposed to  $r$  times (in [3], for example).

Our recursive process relies on an analytic inequality that is simple to describe. Suppose that  $1 \leq r < R$ , and that  $G(\alpha_1, \dots, \alpha_r)$  and  $F(\alpha_1, \dots, \alpha_R)$  are exponential sums, and consider the integral

$$I(F, G) = \int_0^1 \dots \int_0^1 G(\alpha_1, \dots, \alpha_r) F(\alpha_1, \dots, \alpha_R) d\alpha_1 \dots d\alpha_R.$$

Then by Schwarz's inequality, one finds that  $I(F, G)^2 \leq I_1 I_2$ , with

$$I_1 = \int_0^1 \dots \int_0^1 \left| \int_0^1 \dots \int_0^1 F(\alpha_1, \dots, \alpha_R) d\alpha_{r+1} \dots d\alpha_R \right|^2 d\alpha_1 \dots d\alpha_r \quad (1.2)$$

and

$$I_2 = \int_0^1 \dots \int_0^1 |G(\alpha_1, \dots, \alpha_r)|^2 d\alpha_1 \dots d\alpha_r. \quad (1.3)$$

In our application of this inequality, the integral  $I(F, G)$  will count the number of solutions of a system of  $R$  linear equations to be solved in integral cubes. We shall take  $r = 1$  and  $G(\alpha_1)$  equal to an exponential sum related to sums of three cubes. Then, the mean square (1.3) is a sixth moment of cubic Weyl sums for which strong bounds are available. Also, on opening the square in  $I_1$ , a Diophantine interpretation of (1.2) with  $2R - r$  equations is induced. It transpires that this procedure can be repeated, achieving a satisfactory bound for  $I(F, G)$  whenever a good bound for the mean square (1.3) is partnered with good control for the high-dimensional iterates of  $I_1$  that arise from the recursion. While inspired by the work of Gowers [9], the procedure sketched here is in principle very flexible. For example, variants may be developed involving higher moments.

This paper is organised as follows. We begin in §2 by describing the linked block matrices underpinning our new mean value estimates. By using an argument motivated by our earlier work [4], we derive strong estimates associated

with Diophantine systems having six times as many variables as equations. Next, in §3, by repeated application of Schwarz's inequality, we transform an initial system of equations into a more complicated system of the type just analysed. Thus, a powerful mean value estimate is obtained that leads in §4 via the circle method to the proof of Theorem 1.1.

Our basic parameter is  $P$ , a sufficiently large positive number. In this paper, implicit constants in Vinogradov's notation  $\ll$  and  $\gg$  may depend on  $s$ ,  $r$  and  $\varepsilon$ , as well as ambient coefficients. Whenever  $\varepsilon$  appears in a statement, either implicitly or explicitly, we assert that the statement holds for each  $\varepsilon > 0$ . We employ the convention that whenever  $G : [0, 1)^k \rightarrow \mathbb{C}$  is integrable, then

$$\oint G(\alpha) d\alpha = \int_{[0,1)^k} G(\alpha) d\alpha.$$

Here and elsewhere, we use vector notation in the natural way. Finally, we write  $e(z)$  for  $e^{2\pi iz}$  and put  $\|\theta\| = \min\{|\theta - m| : m \in \mathbb{Z}\}$ .

## 2. AUXILIARY EQUATIONS

We begin by defining a strong form of non-singularity satisfied by almost all coefficient matrices. We refer to an  $r \times s$  matrix  $A$  as *highly non-singular* when any subset of at most  $r$  columns of  $A$  is linearly independent. For example, the matrix

$$B = \begin{pmatrix} 7 & 1 & 4 & 8 & 8 & 4 & 9 & 8 & 1 \\ 7 & 5 & 6 & 3 & 3 & 7 & 1 & 7 & 8 \\ 9 & 4 & 5 & 7 & 1 & 6 & 5 & 3 & 6 \\ 6 & 3 & 3 & 8 & 8 & 6 & 9 & 9 & 3 \end{pmatrix}$$

is highly non-singular, as the reader may care to check.

**Lemma 2.1.** *Suppose that the matrix  $A$  is highly non-singular. Then the submatrix obtained by deleting a column is highly non-singular. Also, if a column of  $A$  contains just one non-zero element, then the submatrix obtained by deleting the column and row containing this element is highly non-singular.*

*Proof.* Both conclusions follow from the definition of highly non-singular.  $\square$

Next we describe linked block matrices critical to our arguments. Even to describe the shape of these matrices takes some effort. When  $n$  is a non-negative integer and  $0 \leq l \leq n$ , consider natural numbers  $r_l, s_l$  and an  $r_l \times s_l$  matrix  $A_l$  having non-zero columns. Let  $\text{diag}(A_0, A_1, \dots, A_n)$  be the conventional diagonal block matrix with the lower right hand corner of  $A_l$  sited at  $(i_l, j_l)$ . For  $1 \leq l \leq n$ , append a row to the top of the matrix  $A_l$ , giving an  $(r_l + 1) \times s_l$  matrix  $B_l$ . Next, consider the matrix  $D = (d_{ij})$  obtained from  $\text{diag}(A_0, \dots, A_n)$  by replacing  $A_l$  by  $B_l$  for  $1 \leq l \leq n$ , with the lower right hand corner of  $B_l$  still sited at  $(i_l, j_l)$ . This new *linked-block* matrix  $D$  should be thought of as a matrix with additional entries by comparison to  $\text{diag}(A_0, \dots, A_n)$ , with the property that adjacent blocks are glued together by a shared row sited at index  $i_l$ , for  $0 \leq l < n$ .

**Definition 2.2.** We say that the linked block matrix  $D$  is *congenial of type*  $(n, r; \rho, u, t)$  when it has the shape described above, and

- (a)  $A_l$  and  $B_l$  are highly non-singular, with  $B_l$  of format  $r \times 3(r-1)$ , for  $1 \leq l \leq n$ ;
- (b)  $A_0$  is a  $\rho \times t$  matrix having the following properties:
  - (i) when  $\rho \geq 2$ , its first  $u$  columns define a subspace of dimension 1 distinct from the  $\rho$ -th coordinate axis;
  - (ii) the matrix of its last  $t-u+1$  columns is highly non-singular;
  - (iii) if  $u \geq 3$ , then  $t \geq 3\rho$ .

As a helping hand to the reader, we illustrate this definition with an example. Thus the matrix<sup>1</sup>

$$\begin{pmatrix} 1 & 3 & 3 & 3 & 7 & 1 & 7 & 8 \\ & 7 & 1 & 6 & 5 & 3 & 6 & \\ & 8 & 8 & 6 & 9 & 9 & 3 & 7 & 1 & 4 & 8 & 8 & 4 & 9 & 8 & 1 \\ & & 7 & 5 & 6 & 3 & 3 & 7 & 1 & 7 & 8 \\ & & 9 & 4 & 5 & 7 & 1 & 6 & 5 & 3 & 6 \\ & & 6 & 3 & 3 & 8 & 8 & 6 & 9 & 9 & 3 & 7 & 1 & 4 & 8 & 8 & 4 & 9 & 8 & 1 \\ & & & 7 & 5 & 6 & 3 & 3 & 7 & 1 & 7 & 8 \\ & & & 9 & 4 & 5 & 7 & 1 & 6 & 5 & 3 & 6 \\ & & & 6 & 3 & 3 & 8 & 8 & 6 & 9 & 9 & 3 & 7 & 1 & 4 & 8 & 8 & 4 & 9 & 8 & 1 \\ & & & & 7 & 5 & 6 & 3 & 3 & 7 & 1 & 7 & 8 \\ & & & & 9 & 4 & 5 & 7 & 1 & 6 & 5 & 3 & 6 \\ & & & & 6 & 3 & 3 & 8 & 8 & 6 & 9 & 9 & 3 \end{pmatrix}$$

is congenial of type  $(3, 4; 3, 2, 8)$ . In terms of the description above, one sees that

$$A_0 = \begin{pmatrix} 1 & 3 & 3 & 3 & 7 & 1 & 7 & 8 \\ & 7 & 1 & 6 & 5 & 3 & 6 \\ & 8 & 8 & 6 & 9 & 9 & 3 \end{pmatrix}$$

and  $B_1 = B_2 = B_3 = B$ , and further  $A_1 = A_2 = A_3 = A$ , with

$$A = \begin{pmatrix} 7 & 5 & 6 & 3 & 3 & 7 & 1 & 7 & 8 \\ 9 & 4 & 5 & 7 & 1 & 6 & 5 & 3 & 6 \\ 6 & 3 & 3 & 8 & 8 & 6 & 9 & 9 & 3 \end{pmatrix}.$$

Some additional remarks are in order to clarify this definition. With an inductive argument in mind, we allow the possibility that  $n = 0$ , in which case the parameter  $r$  plays no role. Note that when  $n \geq 1$ , the definition is non-empty only when  $r \geq 2$ . In preparation for our inductive argument, once again, we allow the possibility that  $A_0$  is the empty matrix formally considered to have format  $1 \times 0$ , and we accommodate this situation by identifying congenial matrices (formally) of type  $(n, r; 1, 0, 0)$  with those of type  $(n-1, r; r, 1, 3r-3)$ . Here, as we shall see, there is no loss of generality in assuming that the columns in the first non-empty block of  $D$  have been permuted in order to ensure that its first column is distinct from the  $r$ -th coordinate axis. When  $t \geq 1$ , we insist that  $u \geq 1$ , consistent with the hypothesis that the last  $t-u+1$  columns of  $A_0$  be highly non-singular. Also, we note that when  $\rho = 1$ , the conditions imposed in the preamble to Definition 2.2 require that  $d_{1j} \neq 0$  for  $1 \leq j \leq t$ , and (b) is then satisfied for all  $1 \leq u \leq t$ . When  $\rho \geq 2$ , meanwhile, the value of  $u$  is uniquely determined by the conditions in (b).

<sup>1</sup>Henceforth we adopt the convention that zero entries in a matrix are left blank.

Our goal in this section is to obtain mean value estimates corresponding to auxiliary equations having congenial coefficient matrices. Let  $D$  be an integral congenial matrix of type  $(n, r; \rho, u, t)$ . Then  $D$  is an  $R \times S$  matrix, where  $S = 3n(r-1) + t$  and  $R = n(r-1) + \rho$ . Define the linear forms  $\gamma_j = \gamma_j(\boldsymbol{\alpha})$  by putting

$$\gamma_j(\boldsymbol{\alpha}) = \sum_{i=1}^R d_{ij} \alpha_i \quad (1 \leq j \leq S),$$

and the Weyl sum

$$f(\alpha) = \sum_{|x| \leq P} e(\alpha x^3).$$

Our main lemma provides an estimate for the mean value

$$I(P; D) = \oint |f(\gamma_1) \dots f(\gamma_S)|^2 d\boldsymbol{\alpha}. \quad (2.1)$$

By considering the underlying Diophantine system, one finds that  $I(P; D)$  is unchanged by elementary row operations on  $D$ , and likewise by permutations of its columns. Thus, in the discussion to come we may always pass to a convenient matrix row equivalent to  $D$ .

When  $\rho, w, u$  and  $t$  are non-negative integers, define

$$\delta(\rho, w) = \begin{cases} 1, & \text{when } w = 3\rho, \\ \max\{0, w - 3\rho\}, & \text{otherwise,} \end{cases} \quad (2.2)$$

and then put

$$\delta^*(\rho, u, t) = \begin{cases} \delta(\rho - 1, t - u) + u - 3, & \text{when } \rho \geq 2 \text{ and } u \geq 3, \\ t - 2, & \text{when } \rho = 1 \text{ and } t \geq 3, \\ \delta(\rho, t), & \text{otherwise.} \end{cases} \quad (2.3)$$

**Lemma 2.3.** *When  $\rho \geq 2$ ,  $u \leq t < 3\rho$  and  $u \leq 2$ , one has*

$$\max\{\delta^*(\rho, u, t-1), \delta^*(\rho-1, u, t-1) - 1\} \leq \delta^*(\rho, u, t). \quad (2.4)$$

*Meanwhile, when  $\rho \geq 2$ ,  $t \geq 3\rho$  and  $2 \leq u \leq t$ , one has*

$$\max\{\delta^*(\rho, \max\{u-2, 1\}, t-2) + 1, \delta^*(\rho-1, 1, t-u) + u - 3\} \leq \delta^*(\rho, u, t). \quad (2.5)$$

*Finally, when  $\rho \geq 3$  and  $t \leq u + \rho - 1$ , one has  $\delta^*(\rho-1, u, t) \leq \delta^*(\rho, u, t)$ .*

*Proof.* We first establish (2.4). By (2.3), the inequality to be confirmed reads

$$\max\{\delta(\rho, t-1), \delta(\rho-1, t-1) - 1\} \leq \delta(\rho, t). \quad (2.6)$$

Since  $t \leq 3\rho - 1$ , we have  $\delta(\rho, t-1) = \delta(\rho, t) = 0$ , and since  $t-1 \leq 3(\rho-1) + 1$ , we have also  $\delta(\rho-1, t-1) \leq 1$ . The desired conclusion (2.6) follows.

Suppose next that  $\rho \geq 2$ ,  $t \geq 3\rho$  and  $t \geq u = 2$ . In such circumstances, the inequality (2.5) to be confirmed reads

$$\max\{\delta(\rho, t-2) + 1, \delta(\rho-1, t-2) - 1\} \leq \delta(\rho, t).$$

By considering the cases  $t \in \{3\rho, 3\rho + 1\}$ ,  $t = 3\rho + 2$ , and  $t \geq 3\rho + 3$ , in turn, the desired conclusion follows directly from (2.2). When instead  $u \in \{3, 4\}$ , the inequality (2.5) reads

$$\max\{\delta(\rho, t - 2) + 1, \delta(\rho - 1, t - u) + u - 3\} \leq \delta(\rho - 1, t - u) + u - 3,$$

and one has only to verify that  $\delta(\rho, t - 2) \leq \delta(\rho - 1, t - u) + u - 4$ . By considering the cases  $t \in \{3\rho, 3\rho + 1\}$ ,  $t = 3\rho + 2$ , and  $t \geq 3\rho + 3$ , in turn, the desired conclusion follows directly from (2.2). Finally, when  $t \geq u \geq 5$ , the inequality (2.5) is trivial. We have now confirmed (2.5) in all cases.

In our proof of the final claim of the lemma, we may assume that  $\rho \geq 3$ . Thus  $\rho + 1 \leq 3(\rho - 1) - 1$  and  $\rho - 1 \leq 3(\rho - 2) - 1$ , and hence  $\delta(\rho - 1, \rho + 1) = 0$  and  $\delta(\rho - 2, \rho - 1) = 0$ . Since  $t \leq u + \rho - 1$ , it follows that when  $u \leq 2$ , one has  $\delta(\rho - 1, t) \leq \delta(\rho - 1, \rho + 1) \leq \delta(\rho, t)$ . Likewise, when  $u \geq 3$  it follows that

$$\delta(\rho - 2, t - u) \leq \delta(\rho - 2, \rho - 1) \leq \delta(\rho - 1, t - u).$$

The desired conclusion now follows in both cases from (2.3), completing the proof of the lemma.  $\square$

For future use we record the elementary inequality

$$|z_1 \dots z_n| \leq |z_1|^n + \dots + |z_n|^n. \quad (2.7)$$

**Lemma 2.4.** *Let  $D$  be an integral congenial matrix of type  $(n, r; \rho, u, t)$ . Then*

$$I(P; D) \ll P^{S + \delta^*(\rho, u, t) + \varepsilon}. \quad (2.8)$$

*Proof.* We proceed by induction. Write  $\widehat{H}_{n,r}^{\rho, u, t}$  to denote the hypothesis that the bound (2.8) holds for all congenial matrices of type  $(n, r; \rho, u, t)$ , and  $H_{n,r}^{\rho, u, t}$  to denote the hypothesis that  $\widehat{H}_{n',r'}^{\rho', u', t'}$  holds for all  $n' \leq n$ ,  $r' \leq r$ ,  $\rho' \leq \rho$ ,  $u' \leq u$ ,  $t' \leq t$ . Our outer induction is on  $n$ , with an inner induction on  $r$ ,  $\rho$ ,  $u$  and  $t$ . The basis for this induction is provided by Hua's Lemma (see [15, Lemma 2.5]). This establishes that

$$\int_0^1 |f(\alpha)|^{2u} d\alpha \ll P^{u + \delta(1, u) + \varepsilon} \quad (u = 1, 2, 4).$$

Thus it follows from Hölder's inequality and the trivial estimate  $|f(\alpha)| \leq 2P + 1$  that, when  $n = 0$  and  $\rho = 1$ , then

$$I(P; D) = \int_0^1 |f(\gamma_1) \dots f(\gamma_u)|^2 d\alpha \ll P^{u + \delta(1, u) + \varepsilon} \quad (u \geq 1), \quad (2.9)$$

and one obtains  $H_{0,r}^{1, u, u}$  for all  $u \geq 1$ . Given a congenial matrix  $D$  of type  $(0, r; \rho, u, u)$  with  $\rho \geq 2$ , meanwhile, one has either  $u \leq 2$  or  $u \geq 3\rho$ . It follows by applying elementary row operations that  $D$  is row equivalent to a matrix  $D'$  whose first row entries are all non-zero. By considering the system of equations underlying  $I(P; D')$ , and discarding every equation except that corresponding to the first row of  $D'$ , one finds that  $I(P; D) \leq I(P; D'')$ , where  $D''$  is a congenial matrix of type  $(0, r; 1, u, u)$ . But  $\delta^*(\rho, u, u) \geq \delta(1, u)$  for

$u \leq 2$  and also for  $u \geq 6$ , and thus we deduce from (2.9) that  $\widehat{H}_{0,r}^{\rho,u,u}$  holds for all natural numbers  $\rho$  and  $u$ .

Our strategy for proving the lemma involves two steps. We confirm below that when  $\rho \geq 2$  and  $t > u$ , one has

$$H_{n,r}^{\rho,u,t-1} \text{ implies } \widehat{H}_{n,r}^{\rho,u,t}. \quad (2.10)$$

Notice that when  $\rho = 1$ , then since  $\delta^*(1, t, t) = \delta^*(1, u, t)$ , there is no loss of generality in supposing that  $t = u$ . Since  $u$  (possibly zero) is the smallest value that  $t$  can assume in a congenial matrix of type  $(n, r; \rho, u, t)$ , it therefore suffices to establish  $\widehat{H}_{n,r}^{\rho,u,u}$  ( $u \geq 1$ ). We show below that when  $\rho \geq 1$ , then

$$(\widehat{H}_{n-1,r}^{r,1,3(r-1)} \text{ and } \widehat{H}_{n,r}^{1,u,u}) \text{ implies } \widehat{H}_{n,r}^{\rho,u,u}. \quad (2.11)$$

Since there is no loss in supposing that a congenial matrix of type  $(n, r; 1, u, u)$  is also of type  $(n-1, r; r, \max\{u, 1\}, 3(r-1) + u)$ , one finds via (2.11) that  $H_{n-1,r}^{r,u+1,3r+u}$  implies  $\widehat{H}_{n,r}^{1,u,u}$ , and hence also  $\widehat{H}_{n,r}^{\rho,u,u}$ . We note in this context that  $\delta^*(1, u, u) = \delta^*(r, \max\{u, 1\}, 3(r-1) + u)$ . In view of (2.10), one sees that whenever  $H_{n-1,r}^{\sigma,v,v}$  holds for all  $\sigma$  and  $v$ , then one has  $\widehat{H}_{n,r}^{\rho,u,u}$  for all  $\rho$  and  $u$ , and hence also  $H_{n,r}^{\rho,u,t}$  for all  $\rho, u$  and  $t$ . We have already established  $\widehat{H}_{0,r}^{\sigma,v,v}$  for all  $\sigma$  and  $v$ , and hence the conclusion of the lemma follows by induction on  $n$ .

We begin by confirming (2.11). Let  $D$  be congenial of type  $(n, r; \rho, u, u)$ , and suppose  $\widehat{H}_{n-1,r}^{r,1,3(r-1)}$  and  $\widehat{H}_{n,r}^{1,u,u}$ . We may suppose that  $\rho \geq 2$ , for otherwise (2.11) is trivial. Since the first  $u$  columns of  $D$  define a subspace of dimension 1 distinct from the  $\rho$ -th coordinate axis, the matrix  $D$  has non-zero entries populating one of its first  $\rho-1$  rows in the first  $u$  columns. The matrix  $D$  is consequently row equivalent to one of separated block form, with one block  $D_0$  of format  $1 \times u$  (trivially) congenial of type  $(0, r; 1, u, u)$ , and the second block  $D_1$  of format  $(R-\rho+1) \times (S-u)$ . There is no loss of generality in supposing  $D_1$  to be congenial of type  $(n-1, r; r, 1, 3(r-1))$ . On considering the underlying Diophantine systems, we therefore find that  $I(P; D) = I(P; D_0)I(P; D_1)$ . We may assume (2.9) and  $\widehat{H}_{n-1,r}^{r,1,3(r-1)}$ . Thus we deduce via (2.2) and (2.3) that

$$\begin{aligned} I(P; D) &\ll P^{u+\delta(1,u)+\varepsilon} \cdot P^{S-u+\delta^*(r,1,3(r-1))+\varepsilon} \\ &= P^{S+\delta(1,u)+\delta(r,3(r-1))+2\varepsilon} = P^{S+\delta(1,u)+2\varepsilon}. \end{aligned}$$

When  $\rho \geq 2$ , the congeniality of  $D$  ensures that either  $u \leq 2$  or  $u \geq 6$ , and hence  $\delta^*(\rho, u, u) = \delta(1, u)$ . Thus  $I(P; D) \ll P^{S+\delta^*(\rho,u,u)+\varepsilon}$ , confirming (2.11).

We now commence the proof of (2.10). Let  $D$  be a matrix of type  $(n, r; \rho, u, t)$  with  $\rho \geq 2$  and  $u < t$ , and suppose  $H_{n,r}^{\rho,u,t-1}$ . Should the first  $\rho-1$  rows of the matrix  $D$  be linearly dependent, then by applying elementary row operations on these rows, we may suppose that  $D$  is congenial with one of these rows zero. Thus  $t-u+1 < \rho$  and  $\rho \geq 3$ , and on deleting this row and applying the final conclusion of Lemma 2.3, it is apparent that (2.8) will be confirmed provided that we establish the bound  $I(P; D) \ll P^{S+\delta^*(\rho-1,u,t)+\varepsilon}$ . Repeated use of this simplification permits us to condition the first  $\rho-1$  rows of  $D$  to be linearly independent. We divide into cases according to whether  $t < 3\rho$  or  $t \geq 3\rho$ .



We first establish (2.10) in the situation where  $t < 3\rho$ . One then has  $u \leq 2$ . We distinguish three cases. When  $t = u + 1$ , it follows from the conditioned congeniality of  $D$  that  $\rho = 2$  or  $3$ . In such circumstances, we say that  $D$  has type I when  $\gamma_t = d_{\rho,t}\alpha_\rho$  with  $d_{\rho,t} \neq 0$ . Note that the conditioned congeniality of  $D$  then implies that  $\rho = 2$ . When  $t = u + 1$  and  $D$  is not of type I, we apply elementary row operations to ensure that  $\gamma_t = d_{1,t}\alpha_1$  with  $d_{1,t} \neq 0$ , and also that  $d_{2,j} \neq 0$  for  $1 \leq j \leq u$ . We describe the resulting matrix as having type II. A conditioned congenial matrix  $D$  not of type I or II we describe as having type III. For such matrices, one has  $\rho \geq 3$  and  $t \geq u + 2$ .

Consider first a matrix  $D$  of type I. By performing elementary row operations, one may suppose that  $\gamma_j = d_{1,j}\alpha_1$ , with  $d_{1,j} \neq 0$ , for  $1 \leq j \leq u$ . The matrix  $D$  is of separated block form, with one block  $D_0$  of format  $1 \times u$  (trivially) congenial of type  $(0, 1; 1, u, u)$ , and the second block  $D_1$  of format  $(R - 1) \times (S - u)$  congenial of type  $(n, r; 1, 1, 1)$ . On considering the underlying Diophantine systems, we find that  $I(P; D) \ll I(P; D_0)I(P; D_1)$ . We may assume (2.9) and  $\widehat{H}_{n,r}^{1,1,1}$ , and thus

$$I(P; D) \ll P^{u+\delta(1,u)+\varepsilon} \cdot P^{S-u+\delta^*(1,1,1)+\varepsilon} = P^{S+2\varepsilon} \ll P^{S+\delta^*(2,u,u+1)+\varepsilon},$$

confirming the estimate (2.8) in this case.

Next consider a matrix  $D$  of type III. The last  $t - u + 1$  columns of  $A_0$  span a linear space of dimension  $\min\{t - u + 1, \rho\} \geq 3$ . Hence, by permuting the last  $t - u$  columns of  $A_0$ , we may suppose that the  $t$ -th column of  $A_0$  is not contained in the linear space generated by the first  $u$  columns and the  $\rho$ -th coordinate vector. By applying elementary row operations, we may arrange that the conditioned matrix  $D$  is congenial with  $\gamma_t = d_{1,t}\alpha_1$  and  $d_{1,t} \neq 0$ . We note that the linear space spanned by the first  $u$  columns of  $D$  is now distinct from both the first and the  $\rho$ -th coordinate axis. Let  $D_0$  denote the matrix obtained from  $D$  by deleting column  $t$ , and let  $D_1$  denote the matrix obtained by instead deleting row 1 and column  $t$ . Lemma 2.1 shows the  $R \times (S - 1)$  matrix  $D_0$  to be congenial of type  $(n, r; \rho, u, t - 1)$ , and the  $(R - 1) \times (S - 1)$  matrix  $D_1$  to be congenial of type  $(n, r; \rho - 1, u, t - 1)$ . Observe that when  $D$  is a matrix of type II, then  $\rho = 2$  or  $3$ , and this same conclusion holds.

We may now consider matrices  $D$  of types II and III together. We have  $\gamma_t = d_{1,t}\alpha_1$  with  $d_{1,t} \neq 0$ . Weyl differencing (see [15, equation (2.6)]) yields

$$|f(\gamma_t)|^2 \ll P + \sum_{0 < |h| \leq 16P^3} c_h e(\gamma_t h),$$

where the integers  $c_h$  satisfy  $c_h = O(|h|^\varepsilon)$ . We therefore find from (2.1) that

$$I(P; D) \ll PT(0) + \sum_{0 < |h| \leq 16P^3} c_h T(h), \quad (2.12)$$

where

$$T(h) = \oint \prod_{\substack{1 \leq i \leq S \\ i \neq t}} |f(\gamma_i)|^2 e(\gamma_t h) d\alpha. \quad (2.13)$$

The contribution of the terms with  $h \neq 0$  in (2.12) is given by

$$\sum_{0 < |h| \leq 16P^3} c_h T(h) \ll P^\varepsilon \oint \prod_{\substack{1 \leq i \leq S \\ i \neq t}} |f(\widehat{\gamma}_i)|^2 d\widehat{\alpha}, \quad (2.14)$$

where  $\widehat{\alpha} = (\alpha_2, \dots, \alpha_R)$  and  $\widehat{\gamma}_m = \gamma_m(0, \alpha_2, \dots, \alpha_R)$ . On considering the underlying Diophantine systems, we discern on the one hand from (2.13) that  $T(0) = I(P; D_0)$ , and on the other that the integral on the right hand side of (2.14) is equal to  $I(P; D_1)$ . Thus

$$I(P; D) \ll PI(P; D_0) + P^\varepsilon I(P; D_1).$$

We may assume  $H_{n,r}^{\rho,u,t-1}$ , and thus Lemma 2.3 yields the estimate

$$I(P; D) \ll P^{S+\delta^*(\rho,u,t-1)+\varepsilon} + P^{S-1+\delta^*(\rho-1,u,t-1)+2\varepsilon} \ll P^{S+\delta^*(\rho,u,t)+2\varepsilon}.$$

This confirms the bound (2.8), and hence (2.10) holds whenever  $t < 3\rho$ .

We turn next to the situation with  $t \geq 3\rho$ . Recall that  $\rho \geq 2$  and  $u \geq 1$ . By relabelling the first  $t$  columns of  $D$ , we may assume without loss that the conditioned congenial matrix  $D$  has the property that, should any one of these columns lie on the  $\rho$ -th coordinate axis, then this is the  $t$ -th column. Then, applying the bound (2.7) within (2.1), one finds that with  $j = 1$  or  $2$ , one has

$$I(P; D) \ll \oint |f(\gamma_j)^4 f(\gamma_3)^2 \dots f(\gamma_S)^2| d\alpha. \quad (2.15)$$

Thus, by symmetry, we may suppose that  $j = 1$  and  $u \geq 2$ . Since the first  $u$  columns of  $D$  lie in a subspace of dimension 1 distinct from the  $\rho$ -th coordinate axis, by applying elementary row operations, we see that there is no loss of generality in assuming that the congenial matrix  $D$  satisfies the condition that  $\gamma_j = d_{1,j}\alpha_1$  with  $d_{1,j} \neq 0$  for  $1 \leq j \leq u$ .

We first examine the situation in which  $\rho \geq 2$ ,  $2 \leq u \leq 4$  and  $t \geq 3\rho \geq 6$ . By Weyl differencing (see [15, equation (2.6)]), one has

$$|f(\gamma_1)|^4 \ll P^3 + P \sum_{0 < |h| \leq 32P^3} b_h e(\gamma_1 h),$$

where the integers  $b_h$  satisfy  $b_h = O(|h|^\varepsilon)$ . We therefore find from (2.15) that

$$I(P; D) \ll P^3 U(0) + P \sum_{0 < |h| \leq 32P^3} b_h U(h), \quad (2.16)$$

where

$$U(h) = \oint |f(\gamma_3) \dots f(\gamma_S)|^2 e(\gamma_1 h) d\alpha. \quad (2.17)$$

The contribution of the terms with  $h \neq 0$  in (2.16) is given by

$$P \sum_{0 < |h| \leq 32P^3} b_h U(h) \ll P^{1+\varepsilon} \oint |f(\widehat{\gamma}_3) \dots f(\widehat{\gamma}_S)|^2 d\widehat{\alpha}. \quad (2.18)$$

Let  $D_0$  now denote the matrix obtained from  $D$  by deleting the first two columns, and let  $D_1$  denote the matrix obtained by instead deleting row 1 and the first  $u$  columns. Since  $t \geq 6$ , Lemma 2.1 shows the  $R \times (S-2)$  matrix  $D_0$

to be congenial of type  $(n, r; \rho, \max\{u-2, 1\}, t-2)$ , and the  $(R-1) \times (S-u)$  matrix  $D_1$  to be congenial of type  $(n, r; \rho-1, 1, t-u)$ . On considering the underlying Diophantine systems, we find on the one hand from (2.17) that  $U(0) = I(P; D_0)$ , and on the other that the integral on the right hand side of (2.18) is equal to  $P^{2u-4}I(P; D_1)$ . Here, we have made use of the fact that  $\hat{\gamma}_m(\alpha) = 0$  for  $3 \leq m \leq u$ . Thus

$$I(P; D) \ll P^3 I(P; D_0) + P^{2u-3+\varepsilon} I(P; D_1).$$

We may assume  $H_{n,r}^{\rho,u,t-1}$ , and thus Lemma 2.3 delivers the estimate

$$\begin{aligned} I(P; D) &\ll P^{S+1+\delta^*(\rho, \max\{u-2, 1\}, t-2)+\varepsilon} + P^{S+u-3+\delta^*(\rho-1, 1, t-u)+2\varepsilon} \\ &\ll P^{S+\delta^*(\rho, u, t)+2\varepsilon}. \end{aligned}$$

Since  $\delta^*(\rho, 1, t) = \delta^*(\rho, 2, t)$ , we obtain (2.8) even when the case  $u = 1$  was simplified to that with  $u = 2$ .

Finally, suppose that  $u \geq 5$ . Recall that  $\gamma_j = d_{1,j}\alpha_1$ , with  $d_{1,j} \neq 0$ , for  $1 \leq j \leq u$ . Let  $D_0$  denote the matrix obtained from  $D$  by deleting all but the first row and all but the first  $u$  columns, and let  $D_1$  denote the matrix obtained by instead deleting the first row and first  $u$  columns. Then the  $1 \times u$  matrix  $D_0$  is (trivially) congenial of type  $(0, 1; 1, u, u)$ , and the  $(R-1) \times (S-u)$  matrix  $D_1$  is congenial of type  $(n, r; \rho-1, 1, t-u)$ . On considering the underlying Diophantine systems and applying the triangle inequality, we find via (2.9) that

$$I(P; D) \ll I(P; D_0)I(P; D_1) \ll P^{2u-3+\varepsilon} I(P; D_1),$$

which as above confirms the estimate (2.8) in this final case. Hence we have completed the proof of (2.10) when  $t \geq 3\rho$ , completing the proof of the inductive step. The conclusion of the lemma now follows.  $\square$

We extract a simple consequence from this lemma for future use.

**Corollary 2.5.** *Let  $r \geq 2$ , suppose that  $D$  is an integral congenial matrix of type  $(n, r; r, 3, 3r)$ , and write  $w = (n+1)r - n$ . Then  $I(P; D) \ll P^{3w+1+\varepsilon}$ .*

*Proof.* We have only to note that  $\delta^*(r, 3, 3r) = \delta(r-1, 3r-3) = 1$ .  $\square$

### 3. COMPLIFICATION

Before describing the process which leads from the basic mean value to the more complicated ones described in the previous section, we introduce some additional Weyl sums. When  $2 \leq R \leq P$ , we put

$$\mathcal{A}(P, R) = \{n \in [-P, P] \cap \mathbb{Z} : p \text{ prime and } p|n \Rightarrow p \leq R\},$$

and then define the exponential sum  $g(\alpha) = g(\alpha; P, R)$  by

$$g(\alpha; P, R) = \sum_{x \in \mathcal{A}(P, R)} e(\alpha x^3).$$

We find it convenient to write  $\tau$  for any positive number satisfying  $\tau^{-1} > 852 + 16\sqrt{2833} = 1703.6\dots$ , and then put  $\xi = \frac{1}{4} - \tau$ .

**Lemma 3.1.** *When  $\eta$  is sufficiently small and  $2 \leq R \leq P^\eta$ , one has*

$$\int_0^1 |g(\alpha; P, R)|^6 d\alpha \ll P^{3+\xi}.$$

*Proof.* The conclusion follows from [17, Theorem 1.2] by considering the underlying Diophantine equations.  $\square$

Next we establish an auxiliary lemma that executes the complication process. Let  $n$  and  $r$  be non-negative integers with  $r \geq 2$ , and write  $R = n(r-1)$  and  $S = 3R$ . Let  $B = (b_{i,j})$  be an integral  $(R+1) \times (S+2)$  matrix, write  $\mathbf{b}_j$  for the column vector  $(b_{i,j})_{1 \leq i \leq R+1}$ , and define  $\mathbf{b}_j^*$  to be the column vector  $(b_{R+2-i,j})_{1 \leq i \leq R+1}$  in which the entries of  $\mathbf{b}_j$  are flipped upside down. Also, define  $\beta_j = \beta_j(\boldsymbol{\alpha})$  by putting

$$\beta_j(\boldsymbol{\alpha}) = \sum_{i=1}^{R+1} b_{ij} \alpha_i \quad (0 \leq j \leq S+1). \quad (3.1)$$

We say that the matrix  $B$  is *bicongenial of type  $(n, r)$*  when (i) the column vectors  $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_S$  and  $\mathbf{b}_{S+1}^*, \mathbf{b}_S^*, \dots, \mathbf{b}_1^*$  both form congenial matrices having type  $(n-1, r; r, 1, 3r-2)$ , and (ii) one has  $\beta_0(\boldsymbol{\alpha}) = b_{1,0}\alpha_1$  and  $\beta_{S+1}(\boldsymbol{\alpha}) = b_{R+1,S+1}\alpha_{R+1}$ . At this point, we introduce the mean value

$$J(P; B) = \oint |g(\beta_0)^3 f(\beta_1)^2 \dots f(\beta_S)^2 g(\beta_{S+1})^3| d\boldsymbol{\alpha}. \quad (3.2)$$

Finally, we fix  $\eta > 0$  to be sufficiently small in the context of Lemma 3.1.

**Lemma 3.2.** *Suppose that  $B$  is an integral bicongenial matrix of type  $(n, r)$ . Then there exists an integral bicongenial matrix  $B^*$  of type  $(2n, r)$  for which*

$$J(P; B) \ll (P^{3+\xi})^{1/2} J(P; B^*)^{1/2}.$$

*Proof.* Define the linear forms  $\beta_j$  as in (3.1). Also, define

$$T(P; B) = \int_0^1 \left( \oint |g(\beta_0)^3 f(\beta_1)^2 \dots f(\beta_S)^2| d\widehat{\boldsymbol{\alpha}}_R \right)^2 d\alpha_{R+1},$$

where  $d\widehat{\boldsymbol{\alpha}}_R$  denotes  $d\alpha_1 \dots d\alpha_R$ . Then Schwarz's inequality leads from (3.2) to the bound

$$J(P; B) \leq \left( \int_0^1 |g(\beta_{S+1})|^6 d\alpha_{R+1} \right)^{1/2} T(P; B)^{1/2}. \quad (3.3)$$

By expanding the square inside the outermost integration, we see that

$$T(P; B) = \oint |g(\beta_0^*)^3 f(\beta_1^*)^2 \dots f(\beta_{2S}^*)^2 g(\beta_{2S+1}^*)^3| d\widehat{\boldsymbol{\alpha}}_{2R+1},$$

where  $\beta_i^* = \beta_i^*(\boldsymbol{\alpha})$  is defined by

$$\beta_i^*(\boldsymbol{\alpha}) = \begin{cases} \beta_i(\alpha_1, \dots, \alpha_{R+1}), & \text{when } 0 \leq i \leq S, \\ \beta_{2S+1-i}(\alpha_{2R+1}, \dots, \alpha_{R+1}), & \text{when } S+1 \leq i \leq 2S+1. \end{cases}$$

The integral  $(2R + 1) \times (2S + 2)$  matrix  $B^* = (b_{ij}^*)$  defining the linear forms  $\beta_0^*, \dots, \beta_{2S+1}^*$  is bicongenial of type  $(2n, r)$ , and one has  $T(P; B) = J(P; B^*)$ . The conclusion of the lemma therefore follows from (3.3) and Lemma 3.1.  $\square$

While Lemma 3.2 bounds  $J(P; B)$  in terms of a mean value almost twice the original dimension, superficially *complicating* the task at hand, the higher dimension in fact *simplifies* the problem of obtaining close to square root cancellation. Hence our use of the term *complication*.

Consider an  $r \times s$  integral matrix  $C = (c_{ij})$ , write  $\mathbf{c}_j$  for the column vector  $(c_{ij})_{1 \leq i \leq r}$ , and put

$$\gamma_j = \sum_{i=1}^r c_{ij} \alpha_i \quad (1 \leq j \leq s). \quad (3.4)$$

Also, when  $s \geq 3$ , write

$$K(P; C) = \oint |g(\gamma_1)g(\gamma_2)g(\gamma_3)f(\gamma_4) \dots f(\gamma_s)|^2 d\boldsymbol{\alpha}. \quad (3.5)$$

**Theorem 3.3.** *Suppose that  $r \geq 2$  and that the  $r \times 3r$  integral matrix  $C$  is highly non-singular. Then  $K(P; C) \ll P^{3r+\xi+\varepsilon}$ .*

*Proof.* Write  $s = 3r$ . Since the  $r \times s$  matrix  $C$  is highly non-singular with  $r \geq 2$ , we may apply elementary row operations to  $C$  in such a manner that  $c_{1,1} \neq 0$ ,  $c_{r,2} \neq 0$ , and  $c_{1,3}c_{r,3} \neq 0$ . On considering the underlying Diophantine system, it is apparent from (3.5) that these operations leave the mean value  $K(P; C)$  unchanged. Next, by applying the elementary relation (2.7) within (3.5), one finds by symmetry that there is no loss in supposing that

$$K(P; C) \ll \oint |g(\gamma_1)^3 f(\gamma_4)^2 \dots f(\gamma_s)^2 g(\gamma_2)^3| d\boldsymbol{\alpha}.$$

By relabelling the linear forms, we infer that  $K(P; C) \ll J(P; B_0)$ , where  $B_0$  is the matrix with columns  $\mathbf{c}_1, \mathbf{c}_4, \mathbf{c}_5, \dots, \mathbf{c}_{s-1}, \mathbf{c}_s, \mathbf{c}_2$ . From here, by applying elementary row operations, which amounts to making a non-singular change of variable within (3.5), we may suppose that  $\gamma_1 = c_{1,1}\alpha_1$  and  $\gamma_2 = c_{r,2}\alpha_r$ . Since the  $r \times s$  matrix  $C$  is highly non-singular, Lemma 2.1 shows that  $B_0$  is bicongenial of type  $(1, r)$ .

We show by induction that for each non-negative integer  $l$ , there exists an integral bicongenial matrix of type  $(2^l, r)$  having the property that

$$K(P; C) \ll (P^{3+\xi})^{1-2^{-l}} J(P; B_l)^{2^{-l}}. \quad (3.6)$$

This bound holds when  $l = 0$  as a trivial consequence of the upper bound  $K(P; C) \ll J(P; B_0)$  just established. Suppose then that the estimate (3.6) holds for  $0 \leq l \leq L$ . By applying Lemma 3.2, we see that there exists an integral bicongenial matrix  $B_{L+1}$  of type  $(2^{L+1}, r)$  having the property that

$$J(P; B_L) \ll (P^{3+\xi})^{1/2} J(P; B_{L+1})^{1/2}.$$

Substituting this estimate into the case  $l = L$  of (3.6), one confirms that (3.6) holds with  $l = L + 1$ . The bound (3.6) therefore follows for all  $l$  by induction.

We now prepare to apply the bound just established. Let  $\delta$  be any small positive number, and choose  $l$  large enough that  $2^{1-l}(1-\xi) < \delta$ . We have shown that an integral bicongenial matrix  $B_l = (b_{ij})$  exists for which (3.6) holds. The matrix  $B_l$  is of format  $(R+1) \times (S+2)$ , where  $R = 2^l(r-1)$  and  $S = 3R$ . Define the linear forms  $\beta_j$  as in (3.1) and recall (3.2). Applying (2.7), invoking symmetry, and considering the underlying Diophantine system, we find that there is no loss in supposing that

$$J(P; B_l) \ll \oint |f(\beta_0)^6 f(\beta_1)^2 \dots f(\beta_S)^2| d\alpha.$$

Let  $D$  be the integral matrix underlying the  $S+3$  forms  $\beta_0, \beta_0, \beta_0, \beta_1, \dots, \beta_S$ . Then  $D$  is congenial of type  $(2^l-1, r; r, 3, 3r)$ , and one has  $J(P; B_l) \ll I(P; D)$ . Substituting the bound  $J(P; B_l) \ll P^{3R+4+\varepsilon}$  that follows from Corollary 2.5 into (3.6), we obtain the estimate

$$K(P; C) \ll (P^{3+\xi})^{1-2^{-l}} (P^{3(2^l(r-1)+1)+1+\varepsilon})^{2^{-l}} \ll P^{3r+\xi+(1-\xi)2^{-l}+\varepsilon}.$$

In view of our assumed upper bound  $2^{1-l}(1-\xi) < \delta$ , one therefore sees that

$$K(P; C) \ll P^{3r+\xi+\frac{1}{2}\delta+\varepsilon} \ll P^{3r+\xi+\delta}.$$

The conclusion of the theorem now follows by taking  $\delta$  sufficiently small.  $\square$

#### 4. THE HARDY-LITTLEWOOD METHOD

In this section we turn to the proof of Theorem 1.1. Let  $(c_{ij})$  denote an integral  $r \times s$  highly non-singular matrix with  $r \geq 2$  and  $s \geq 6r+1$ . We define the linear forms  $\gamma_j = \gamma_j(\alpha)$  as in (3.4), and for concision put  $g_j = g(\gamma_j(\alpha))$  and  $f_j = f(\gamma_j(\alpha))$ . When  $\mathfrak{B} \subseteq [0, 1]^r$  is measurable, we then define

$$N(P; \mathfrak{B}) = \int_{\mathfrak{B}} g_1 \dots g_6 f_7 \dots f_s d\alpha.$$

By orthogonality, it follows from this definition that  $N(P; [0, 1]^r)$  counts the number of integral solutions of the system (1.1) with  $x_1, \dots, x_6 \in \mathcal{A}(P, R)$  and  $x_7, \dots, x_s \in [-P, P]$ . In this section we prove the lower bound  $N(P; [0, 1]^r) \gg P^{s-3r}$ , subject to the hypothesis that the system (1.1) has non-zero  $p$ -adic solutions for all primes  $p$ . The conclusion of Theorem 1.1 then follows.

In pursuit of the above objective, we apply the Hardy-Littlewood method. Let  $\mathfrak{M}$  denote the union of the intervals

$$\mathfrak{M}(q, a) = \{\alpha \in [0, 1) : |q\alpha - a| \leq (6P^2)^{-1}\},$$

with  $0 \leq a \leq q \leq P$  and  $(a, q) = 1$ , and let  $\mathfrak{m} = [0, 1) \setminus \mathfrak{M}$ . In addition, write  $L = \log P$ , denote by  $\mathfrak{N}$  the union of the intervals

$$\mathfrak{N}(q, a) = \{\alpha \in [0, 1) : |q\alpha - a| \leq LP^{-3}\},$$

with  $0 \leq a \leq q \leq L$  and  $(a, q) = 1$ , and put  $\mathfrak{n} = [0, 1) \setminus \mathfrak{N}$ . We summarise some useful estimates in this context in the form of a lemma.

**Lemma 4.1.** *One has*

$$\int_{\mathfrak{M} \setminus \mathfrak{N}} |f(\alpha)|^5 d\alpha \ll P^2 L^{\varepsilon-1/3} \quad \text{and} \quad \int_0^1 |f(\alpha)|^8 d\alpha \ll P^5.$$

*Proof.* The first estimate follows as a special case of [14, Lemma 5.1], and the second is immediate from [13, Theorem 2], by orthogonality.  $\square$

Next we introduce a multi-dimensional set of arcs. Let  $Q = L^{10r}$ , and define the narrow set of major arcs  $\mathfrak{P}$  to be the union of the boxes

$$\mathfrak{P}(q, \mathbf{a}) = \{\alpha \in [0, 1)^r : |\alpha_i - a_i/q| \leq QP^{-3} \ (1 \leq i \leq r)\},$$

with  $0 \leq a_i \leq q \leq Q$  ( $1 \leq i \leq r$ ) and  $(a_1, \dots, a_r, q) = 1$ .

**Lemma 4.2.** *Suppose that the system (1.1) admits non-zero  $p$ -adic solutions for each prime number  $p$ . Then one has  $N(P; \mathfrak{P}) \gg P^{s-3r}$ .*

*Proof.* We begin by defining the auxiliary functions

$$S(q, \mathbf{a}) = \sum_{r=1}^q e(ar^3/q) \quad \text{and} \quad v(\beta) = \int_{-P}^P e(\beta\gamma^3) d\gamma.$$

For  $1 \leq j \leq s$ , put  $S_j(q, \mathbf{a}) = S(q, \gamma_j(\mathbf{a}))$  and  $v_j(\beta) = v(\gamma_j(\beta))$ , and define

$$A(q) = \sum_{\substack{a_1=1 \\ (q, a_1, \dots, a_r)=1}}^q \cdots \sum_{\substack{a_r=1 \\ (q, a_1, \dots, a_r)=1}}^q q^{-s} \prod_{j=1}^s S_j(q, \mathbf{a}) \quad \text{and} \quad V(\beta) = \prod_{j=1}^s v_j(\beta). \quad (4.1)$$

Finally, write  $\mathcal{B}(X)$  for  $[-XP^{-3}, XP^{-3}]^r$ , and define

$$\mathfrak{J}(X) = \int_{\mathcal{B}(X)} V(\beta) d\beta \quad \text{and} \quad \mathfrak{S}(X) = \sum_{1 \leq q \leq X} A(q).$$

We prove first that there exists a positive constant  $C$  with the property that

$$N(P; \mathfrak{P}) - C\mathfrak{S}(Q)\mathfrak{J}(Q) \ll P^{s-3r}L^{-1}. \quad (4.2)$$

It follows from [16, Lemma 8.5] (see also [14, Lemma 5.4]) that there exists a positive constant  $c = c(\eta)$  such that whenever  $\alpha \in \mathfrak{P}(q, \mathbf{a}) \subseteq \mathfrak{P}$ , then

$$g(\gamma_j(\alpha)) - cq^{-1}S_j(q, \mathbf{a})v_j(\alpha - \mathbf{a}/q) \ll P(\log P)^{-1/2}.$$

Under the same constraints on  $\alpha$ , one finds from [15, Theorem 4.1] that

$$f(\gamma_j(\alpha)) - q^{-1}S_j(q, \mathbf{a})v_j(\alpha - \mathbf{a}/q) \ll \log P.$$

Thus, whenever  $\alpha \in \mathfrak{P}(q, \mathbf{a}) \subseteq \mathfrak{P}$ , one has

$$g_1 \cdots g_s f_7 \cdots f_s - c^6 q^{-s} \prod_{j=1}^s S_j(q, \mathbf{a})v_j(\alpha - \mathbf{a}/q) \ll P^s(\log P)^{-1/2}.$$

The measure of the major arcs  $\mathfrak{P}$  is  $O(Q^{2r+1}P^{-3r})$ , so that on integrating over  $\mathfrak{P}$ , we confirm the relation (4.2) with  $C = c^6$ .

We next discuss the singular integral  $\mathfrak{J}(Q)$ . By applying (2.7), we find that

$$V(\boldsymbol{\beta}) \ll \sum_{1 \leq j_1 < \dots < j_r \leq s} |v_{j_1}(\boldsymbol{\beta}) \dots v_{j_r}(\boldsymbol{\beta})|^{s/r}. \quad (4.3)$$

Recall from [15, Theorem 7.3] that  $v(\beta) \ll P(1 + P^3|\beta|)^{-1/3}$ . Since  $(c_{ij})$  is highly non-singular and  $s \geq 6r + 1$ , a change of variables reveals that  $V(\boldsymbol{\beta})$  is integrable, that the limit  $\mathfrak{J} = \lim_{X \rightarrow \infty} \mathfrak{J}(X)$  exists, and that  $\mathfrak{J} \ll P^{s-3r}$ . Write  $\widehat{\mathcal{B}}(X) = \mathbb{R}^r \setminus \mathcal{B}(X)$ . Then by applying (4.3), we discern that there are distinct indices  $j_1, \dots, j_r$  such that

$$\mathfrak{J} - \mathfrak{J}(X) = \int_{\widehat{\mathcal{B}}(X)} V(\boldsymbol{\beta}) d\boldsymbol{\beta} \ll \int_{\widehat{\mathcal{B}}(X)} |v_{j_1}(\boldsymbol{\beta}) \dots v_{j_r}(\boldsymbol{\beta})|^{s/r} d\boldsymbol{\beta}.$$

The linear independence of the  $\gamma_j$  ensures that whenever  $\boldsymbol{\beta} \in \widehat{\mathcal{B}}(X)$ , then for some index  $l$  with  $1 \leq l \leq r$ , one has  $|\gamma_{j_l}(\boldsymbol{\beta})| > X^{1/2}P^{-3}$ . Consequently, the hypothesis  $s \geq 6r + 1$  again ensures via a change of variables that

$$\begin{aligned} \mathfrak{J} - \mathfrak{J}(X) &\ll \left( \sup_{\boldsymbol{\beta} \in \widehat{\mathcal{B}}(X)} |v_{j_1}(\boldsymbol{\beta}) \dots v_{j_r}(\boldsymbol{\beta})| \right) \int_{\widehat{\mathcal{B}}(X)} |v_{j_1}(\boldsymbol{\beta}) \dots v_{j_r}(\boldsymbol{\beta})|^{(s-r)/r} d\boldsymbol{\beta} \\ &\ll P^s X^{-1/6} \int_{\mathbb{R}^r} \prod_{j=1}^r (1 + P^3|\theta_j|)^{-(s-r)/(3r)} d\boldsymbol{\theta} \ll P^{s-3r} X^{-1/6}. \end{aligned}$$

The system of equations (1.1) possesses a non-zero real solution in  $[-1, 1]^s$ , and this must be non-singular since  $(c_{ij})$  is highly non-singular. An application of Fourier's integral formula (see [6, Chapter 4] and [8, Lemma 30]) therefore leads to the lower bound  $\mathfrak{J} \gg P^{s-3r}$ . Thus we may conclude that

$$\mathfrak{J}(Q) \gg P^{s-3r} + O(P^{s-3r}Q^{-1/6}) \gg P^{s-3r}. \quad (4.4)$$

We turn next to the singular series  $\mathfrak{S}(Q)$ . It follows from [15, Theorem 4.2] that whenever  $(q, a) = 1$ , one has  $S(q, a) \ll q^{2/3}$ . Given a summand  $\mathbf{a}$  in the formula for  $A(q)$  provided in (4.1), write  $h_j = (q, \gamma_j(\mathbf{a}))$ . Then we find that

$$A(q) \ll \sum_{\substack{a_1=1 \\ (q, a_1, \dots, a_r)=1}}^q \dots \sum_{a_r=1}^q q^{-s/3} (h_1 \dots h_s)^{1/3}.$$

By hypothesis, we have  $s/(3r) \geq 2 + 1/(3r)$ . The proof of [8, Lemma 23] is therefore easily modified to show that  $A(q) \ll q^{-1-1/(6r)}$ . Thus, the series  $\mathfrak{S} = \lim_{X \rightarrow \infty} \mathfrak{S}(X)$  is absolutely convergent and

$$\mathfrak{S} - \mathfrak{S}(Q) \ll \sum_{q > Q} q^{-1-1/(6r)} \ll Q^{-1/(6r)} \ll L^{-1}.$$

The system (1.1) has non-zero  $p$ -adic solutions for each prime  $p$ , and these are non-singular since  $(c_{ij})$  is highly non-singular. A modification of the proof of [8, Lemma 31] therefore shows that  $\mathfrak{S} > 0$ , whence  $\mathfrak{S}(Q) = \mathfrak{S} + O(L^{-1}) > 0$ . The proof of the lemma is completed by recalling (4.4) and substituting into (4.2) to obtain the bound  $N(P; \mathfrak{P}) \gg P^{s-3r} + O(P^{s-3r}L^{-1})$ .  $\square$



In order to prune a wide set of major arcs down to the narrow set  $\mathfrak{P}$  just considered, we introduce the auxiliary sets of arcs

$$\mathfrak{M}_j = \{\alpha \in [0, 1)^r : \gamma_j(\alpha) \in \mathfrak{M} + \mathbb{Z}\},$$

and we put  $\mathfrak{V} = \mathfrak{M}_7 \cap \mathfrak{M}_8 \cap \dots \cap \mathfrak{M}_s$ . In addition, we define  $\mathfrak{m}_j = [0, 1)^r \setminus \mathfrak{M}_j$  ( $7 \leq j \leq s$ ), and write  $\mathfrak{v} = [0, 1)^r \setminus \mathfrak{V}$ . Finally, for any positive integer  $n$ , when  $\omega \in [1, s]^n$ , we define

$$\mathfrak{K}_\omega = \{\alpha \in \mathfrak{V} \setminus \mathfrak{P} : \gamma_{\omega_m}(\alpha) \in \mathfrak{n} + \mathbb{Z} \ (1 \leq m \leq n)\}.$$

**Lemma 4.3.** *One has  $N(P; \mathfrak{V} \setminus \mathfrak{P}) \ll P^{s-3r} L^{-1/4}$ .*

*Proof.* Let  $\alpha \in \mathfrak{V} \setminus \mathfrak{P}$ , and suppose temporarily that  $\gamma_{j_m} \in \mathfrak{n} + \mathbb{Z}$  for  $r$  distinct indices  $j_m \in [7, s]$ . For each  $m$  there is a natural number  $q_m \leq L$  having the property that  $\|q_m \gamma_{j_m}\| \leq LP^{-3}$ . With  $q = q_1 \dots q_r$ , one has  $q \leq L^r$  and  $\|q \gamma_{j_m}\| \leq L^r P^{-3}$ . Next eliminating between  $\gamma_{j_1}, \dots, \gamma_{j_r}$  in order to isolate  $\alpha_1, \dots, \alpha_r$ , one finds that there is a positive integer  $\kappa$ , depending at most on  $(c_{ij})$ , such that  $\|\kappa q \alpha_l\| \leq L^{r+1} P^{-3}$  ( $1 \leq l \leq r$ ). Since  $\kappa q \leq L^{r+1}$ , it follows that  $\alpha \in \mathfrak{P}$ , yielding a contradiction to our hypothesis that  $\alpha \in \mathfrak{V} \setminus \mathfrak{P}$ . Thus  $\gamma_\nu(\alpha) \in \mathfrak{n} + \mathbb{Z}$  for at least  $s - 6 - r \geq 5(r - 1)$  of the suffices  $\nu$  with  $7 \leq \nu \leq s$ . Then for some tuple  $\nu = (\nu_1, \dots, \nu_{5r-5})$  of distinct integers  $\nu_m \in [7, s]$ , one has

$$N(P; \mathfrak{V} \setminus \mathfrak{P}) \ll \int_{\mathfrak{K}_\nu} |g_1 \dots g_6 f_7 \dots f_s| d\alpha.$$

By symmetry, we may suppose that  $\nu = (9, \dots, 5r + 3)$ . Let  $k_l$  denote  $g_l$  when  $1 \leq l \leq 6$ , and  $f_l$  when  $l = 7, 8$ . Then combining (2.7) with a trivial estimate for  $|f(\alpha)|$ , one finds that for some tuple  $(\sigma_1, \dots, \sigma_{r-1})$  of distinct integers  $\sigma_m \in [9, 5r + 3]$ , and some integer  $l$  with  $1 \leq l \leq 8$ , one has

$$N(P; \mathfrak{V} \setminus \mathfrak{P}) \ll P^{s-5r-3} \int_{\mathfrak{K}_\sigma} |k_l^8 f_{\sigma_1}^5 \dots f_{\sigma_{r-1}}^5| d\alpha.$$

By changing variables, considering the underlying Diophantine equations, and applying Lemma 4.1, we deduce that

$$\begin{aligned} N(P; \mathfrak{V} \setminus \mathfrak{P}) &\ll P^{s-5r-3} \left( \int_0^1 |f(\alpha)|^8 d\alpha \right) \left( \int_{\mathfrak{M} \setminus \mathfrak{N}} |f(\alpha)|^5 d\alpha \right)^{r-1} \\ &\ll P^{s-5r-3} (P^5) (P^2 L^{\varepsilon-1/3})^{r-1} \ll P^{s-3r} L^{-1/4}, \end{aligned}$$

and the proof of the lemma is complete.  $\square$

**Lemma 4.4.** *There is a positive number  $\delta$  such that  $N(P; \mathfrak{v}) \ll P^{s-3r-\delta}$ .*

*Proof.* If  $\alpha \in \mathfrak{v}$ , then for some index  $j$  with  $7 \leq j \leq s$ , one has  $\gamma_j(\alpha) \notin \mathfrak{M} + \mathbb{Z}$ , and so  $\alpha \in \mathfrak{m}_j$ . Thus, combining (2.7) with a trivial estimate for  $|f(\alpha)|$ , we find that for some suffix  $j \in [7, s]$ , and some tuple  $(j_1, \dots, j_{3r})$  with

$$1 \leq j_1 < j_2 < j_3 \leq 6 < j_4 < \dots < j_{3r} \leq s,$$

one has

$$N(P; \mathfrak{v}) \ll P^{s-6r-1} \sup_{\alpha \in \mathfrak{m}_j} |f(\gamma_j(\alpha))| \oint |g_{j_1} g_{j_2} g_{j_3} f_{j_4} \dots f_{j_{3r}}|^2 d\alpha. \quad (4.5)$$

The matrix underlying the linear forms  $\gamma_{j_1}, \dots, \gamma_{j_{3r}}$  is highly non-singular, and so we may apply Theorem 3.3 to estimate the integral on the right hand side of (4.5). Moreover, by Weyl's inequality (see [15, Lemma 2.4]), one has

$$\sup_{\alpha \in \mathfrak{m}_j} |f(\gamma_j(\alpha))| \leq \sup_{\beta \in \mathfrak{m}} |f(\beta)| \ll P^{3/4+\varepsilon}.$$

We therefore conclude that for some positive number  $\delta$ , one has

$$N(P; \mathfrak{v}) \ll P^{s-6r-1} (P^{3/4+\varepsilon}) (P^{3r+\xi+\varepsilon}) \ll P^{s-3r-\delta}.$$

This completes the proof of the lemma.  $\square$

By combining Lemmata 4.2, 4.3 and 4.4, we infer that whenever the system (1.1) possesses a non-zero  $p$ -adic solution, one has

$$\begin{aligned} N(P) &= N(P; \mathfrak{P}) + N(P; \mathfrak{V} \setminus \mathfrak{P}) + N(P; \mathfrak{v}) \\ &\gg P^{s-3r} + O(P^{s-3r} L^{-1/4} + P^{s-3r-\delta}) \gg P^{s-3r}. \end{aligned}$$

This completes our proof of Theorem 1.1.

We remark that the condition in Theorem 1.1 that  $(c_{ij})$  be highly non-singular can certainly be relaxed. Let us refer to the number of columns lying in a given one dimensional subspace of the column space of  $(c_{ij})$  as the *multiplicity* of that subspace. The discussion of §§2 and 3 would suffer no ill consequences were  $(c_{ij})$  to satisfy the condition that the maximum multiplicity be 2. In order to see this, one has simply to note that in such circumstances, the mean value estimates relevant to the application of the Hardy-Littlewood method can be related, via Hölder's inequality, to mean values of the shape (3.5). We note in this context that the matrix  $(c_{ij})$  occurring in Theorem 1.1 is of course different from that occurring in Theorem 3.3. With rather greater effort in a more cumbersome argument, this maximum multiplicity 2 could be increased to 3, and even several multiplicities of 4 can be tolerated. This and further refinements are topics that we intend to pursue on a future occasion.

**Acknowledgement.** The authors are grateful to the referees for the extreme care taken in reviewing this paper, and in particular for numerous suggestions which have clarified our exposition and prompted significant corrections.

## REFERENCES

- [1] O. D. Atkinson, J. Brüdern and R. J. Cook, *Simultaneous additive congruences to a large prime modulus*, *Mathematika* **39** (1992), 1–9.
- [2] R. C. Baker, *Diagonal cubic equations III*, *Proc. London Math. Soc.* (3) **58** (1989), 495–518.
- [3] J. Brüdern and R. J. Cook, *On simultaneous diagonal equations and inequalities*, *Acta Arith.* **62** (1992), 125–149.
- [4] J. Brüdern and T. D. Wooley, *Hua's Lemma and simultaneous diagonal equations*, *Bull. London Math. Soc.* **34** (2002), 279–283.
- [5] J. Brüdern and T. D. Wooley, *The Hasse principle for pairs of diagonal cubic forms*, *Ann. of Math.* (2) **166** (2007), 865–895.
- [6] H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, 2nd edition, Cambridge University Press, Cambridge, 2005.

- [7] H. Davenport and D. J. Lewis, *Cubic equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A **261** (1966), 97–136.
- [8] H. Davenport and D. J. Lewis, *Simultaneous equations of additive type*, Philos. Trans. Roy. Soc. London Ser. A **264** (1969), 557–595.
- [9] W. T. Gowers, *A new proof of Szemerédi’s theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.
- [10] D. R. Heath-Brown, *The circle method and diagonal cubic forms*, Phil. Trans. Roy. Soc. London Ser. A **356** (1998), 673–699.
- [11] C. Hooley, *On Waring’s problem*, Acta Math. **157** (1986), 49–97.
- [12] C. Hooley, *On Hypothesis  $K^*$  in Waring’s problem*, Sieve methods, exponential sums, and their applications in number theory (Cardiff, 1995), pp. 175–185, Cambridge University Press, 1997.
- [13] R. C. Vaughan, *On Waring’s problem for cubes*, J. Reine Angew. Math. **365** (1986), 122–170.
- [14] R. C. Vaughan, *A new iterative method in Waring’s problem*, Acta Math. **162** (1989), 1–71.
- [15] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd edition, Cambridge University Press, Cambridge, 1997.
- [16] T. D. Wooley, *On simultaneous additive equations, II*, J. Reine Angew. Math. **419** (1991), 141–198.
- [17] T. D. Wooley, *Sums of three cubes*, Mathematika **47** (2000), 53–61.

MATHEMATISCHES INSTITUT, BUNSENSTRASSE 3–5, D-37073 GÖTTINGEN, GERMANY  
*E-mail address:* bruedern@uni-math.gwdg.de

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, UNIVERSITY WALK, CLIFTON,  
BRISTOL BS8 1TW, UNITED KINGDOM  
*E-mail address:* matdw@bristol.ac.uk